

71
12

RECEIVED

APR 07 2003

RECEIVED

DOCK 1-1-1



TECH CENTER 1600/2900 1600

RAW SEQUENCE LISTING

DATE: 02/25/2003

PATENT APPLICATION: US/09/576,778B

TIME: 11:45:52

Input Set : N:\Crif4\02242003\I576778C.raw

Output Set: N:\CRF4\02252003\I576778B.raw

1 <110> APPLICANT: Schulein, Martin
2 Bjornvad, Mads
3 <120> TITLE OF INVENTION: Family 9 Endo-Beta-1, 4-Glucanases
4 <130> FILE REFERENCE: 5843.200-US
C--> 5 <140> CURRENT APPLICATION NUMBER: US/09/576,778B
6 <141> CURRENT FILING DATE: 2000-05-23
7 <160> NUMBER OF SEQ ID NOS: 10
8 <170> SOFTWARE: PatentIn version 3.1
10 <210> SEQ ID NO: 1
11 <211> LENGTH: 1941
12 <212> TYPE: DNA
13 <213> ORGANISM: Bacillus licheniformis
14 <220> FEATURE:
15 <221> NAME/KEY: CDS
16 <222> LOCATION: (1)..(1938)
17 <223> OTHER INFORMATION:
18 <220> FEATURE:
19 <221> NAME/KEY: sig_peptide
20 <222> LOCATION: (1)..(87)
21 <223> OTHER INFORMATION:
W--> 22 <220>
23 <221> NAME/KEY: mat_peptide
24 <222> LOCATION: (88)..()
25 <223> OTHER INFORMATION:

ENTERED

W--> 26 <400> 1
27 atg aaa gcg ctt tgt ttg gcw ctt tta gtg atc ttc tct atg agc ata 48
28 Met Lys Ala Leu Cys Leu Ala Leu Leu Val Ile Phe Ser Met Ser Ile
29 -25 -20 -15
30 gcg tcg ttt tca gaa aag acc cgt gca gct tct gct gaa gaa tat cct 96
31 Ala Ser Phe Ser Glu Lys Thr Arg Ala Ala Ser Ala Glu Glu Tyr Pro
32 -10 -5 -1 1
33 cat aat tat gct gaa ctg ctg caa aag tct ttg tta ttt tat gaa gca 144
34 His Asn Tyr Ala Glu Leu Leu Gln Lys Ser Leu Leu Phe Tyr Glu Ala
35 5 10 15
36 cag cgc tcg gga aga ctt ccg gaa aac agc cgg ctg aat tgg aga gga 192
37 Gln Arg Ser Gly Arg Leu Pro Glu Asn Ser Arg Leu Asn Trp Arg Gly
38 20 25 30 35
39 gac tcc ggg ctt gag gac gga aaa gac gtt ggc ctc gat tta acg gga 240
40 Asp Ser Gly Leu Glu Asp Gly Lys Asp Val Gly Leu Asp Leu Thr Gly
41 40 45 50
42 ggg tgg tat gat gcc ggc gac cac gtg aag ttc ggt ctg ccg atg gct 288
43 Gly Trp Tyr Asp Ala Gly Asp His Val Lys Phe Gly Leu Pro Met Ala
44 55 60 65

RAW SEQUENCE LISTING

DATE: 02/25/2003

PATENT APPLICATION: US/09/576,778B

TIME: 11:45:52

Input Set : N:\Crf4\02242003\I576778C.raw

Output Set: N:\CRF4\02252003\I576778B.raw

45	tat tct gcc gca atc ctg tca tgg tgg gtc tat gag tac cga gat gcc	336
46	Tyr Ser Ala Ala Ile Leu Ser Trp Ser Val Tyr Glu Tyr Arg Asp Ala	
47	70 75 80	
48	tac aaa gaa tgg ggt cag ctt gat gcg gcg ctg gac aat att aaa tgg	384
49	Tyr Lys Glu Ser Gly Gln Leu Asp Ala Ala Leu Asp Asn Ile Lys Trp	
50	85 90 95	
51	gcg aca gac tac ttt ctt aaa gcc cat acg gct cct tat gaa ttg tgg	432
52	Ala Thr Asp Tyr Phe Leu Lys Ala His Thr Ala Pro Tyr Glu Leu Trp	
53	100 105 110 115	
54	ggc caa gtc gga aat ggc gct cta gac cac gca tgg tgg ggg ccg gcc	480
55	Gly Gln Val Gly Asn Gly Ala Leu Asp His Ala Trp Trp Gly Pro Ala	
56	120 125 130	
57	gaa gta atg ccg atg aag cgc cct gcc tat aag atc gat gcc ggc tgt	528
58	Glu Val Met Pro Met Lys Arg Pro Ala Tyr Lys Ile Asp Ala Gly Cys	
59	135 140 145	
60	ccg ggg tca gac ctt gct ggt ggt aca gcc gca gcg cta gca tca gca	576
61	Pro Gly Ser Asp Leu Ala Gly Gly Thr Ala Ala Ala Leu Ala Ser Ala	
62	150 155 160	
63	tca att att ttc aag ccg aca gat tct tct tac tct gaa aaa tta ctg	624
64	Ser Ile Ile Phe Lys Pro Thr Asp Ser Ser Tyr Ser Glu Lys Leu Leu	
65	165 170 175	
66	gct cat gcc aag caa ttg tat gat ttt gcc gac cgc tac cgc ggc aaa	672
67	Ala His Ala Lys Gln Leu Tyr Asp Phe Ala Asp Arg Tyr Arg Gly Lys	
68	180 185 190 195	
69	tat tca gac tgc att aca gac gca cag caa tat tat aat tgg tgg agc	720
70	Tyr Ser Asp Cys Ile Thr Asp Ala Gln Gln Tyr Tyr Asn Ser Trp Ser	
71	200 205 210	
72	ggg tat aaa gat gaa ctg aca tgg gga gct gtc tgg ctc tac ttg gca	768
73	Gly Tyr Lys Asp Glu Leu Thr Trp Gly Ala Val Trp Leu Tyr Leu Ala	
74	215 220 225	
75	aca gaa gaa caa caa tat ttg gat aaa gcc ctt gct tgg gtc tca gat	816
76	Thr Glu Glu Gln Gln Tyr Leu Asp Lys Ala Leu Ala Ser Val Ser Asp	
77	230 235 240	
78	tgg ggc gat ccc gca aac tgg cct tac cgc tgg acg ctt tcc tgg gat	864
79	Trp Gly Asp Pro Ala Asn Trp Pro Tyr Arg Trp Thr Leu Ser Trp Asp	
80	245 250 255	
81	gac gtc act tac gga gca cag ctg ctg ctc gct cgt ctg aca aac gat	912
82	Asp Val Thr Tyr Gly Ala Gln Leu Leu Leu Ala Arg Leu Thr Asn Asp	
83	260 265 270 275	
84	tcc cgt ttt gtc aaa tct gtc gaa cgc aat ctt gat tat tgg tgg aca	960
85	Ser Arg Phe Val Lys Ser Val Glu Arg Asn Leu Asp Tyr Trp Ser Thr	
86	280 285 290	
87	ggc tac agt cat aat gga agc ata gaa cgg atc acg tat acg ccg ggc	1008
88	Gly Tyr Ser His Asn Gly Ser Ile Glu Arg Ile Thr Tyr Thr Pro Gly	
89	295 300 305	
90	ggt ttg gcc tgg ctt gag cag tgg gga tca ttg cga tac gct tgg aat	1056
91	Gly Leu Ala Trp Leu Glu Gln Trp Gly Ser Leu Arg Tyr Ala Ser Asn	
92	310 315 320	
93	gcc gct ttt ctc gct ttc gtt tat tcc gat tgg gtc gat aca gaa aaa	1104

RAW SEQUENCE LISTING

DATE: 02/25/2003

PATENT APPLICATION: US/09/576,778B

TIME: 11:45:52

Input Set : N:\Crf4\02242003\I576778C.raw

Output Set: N:\CRF4\02252003\I576778B.raw

94	Ala Ala Phe Leu Ala Phe Val Tyr Ser Asp Trp Val Asp Thr Glu Lys	
95	325 330 335	
96	gcg aaa aga tat cgg gat ttt gct gtt cgg caa acg gag tat atg cta	1152
97	Ala Lys Arg Tyr Arg Asp Phe Ala Val Arg Gln Thr Glu Tyr Met Leu	
98	340 345 350 355	
99	gga gat aat ccg cag cag cga agc ttt gtc gtt gga tac ggt aaa aat	1200
100	Gly Asp Asn Pro Gln Gln Arg Ser Phe Val Val Gly Tyr Gly Lys Asn	
101	360 365 370	
102	ccg ccg aaa cat ccg cat cac cgt aca gca cac ggt tca tgg gcc aat	1248
103	Pro Pro Lys His Pro His His Arg Thr Ala His Gly Ser Trp Ala Asn	
104	375 380 385	
105	cag atg aat gtg cct gaa aac cat cgc cat acc cta tac ggc gca tta	1296
106	Gln Met Asn Val Pro Glu Asn His Arg His Thr Leu Tyr Gly Ala Leu	
107	390 395 400	
108	gtc ggc ggt ccg gga agg gac gat tcg tac cga gat gac ata aca gat	1344
109	Val Gly Gly Pro Gly Arg Asp Ser Tyr Arg Asp Asp Ile Thr Asp	
110	405 410 415	
111	tat gcg tca aac gaa gtt gcg atc gat tat aat gcc gct ttt acc ggc	1392
112	Tyr Ala Ser Asn Glu Val Ala Ile Asp Tyr Asn Ala Ala Phe Thr Gly	
113	420 425 430 435	
114	aac gta gcg aaa atg ttt cag ctg ttc ggg aaa ggc cat gtt ccg ctg	1440
115	Asn Val Ala Lys Met Phe Gln Leu Phe Gly Lys Gly His Val Pro Leu	
116	440 445 450	
117	cct gat ttt ccg gag aag gaa aca cct gag gac gaa tat ttt gca gag	1488
118	Pro Asp Phe Pro Glu Lys Glu Thr Pro Glu Asp Glu Tyr Phe Ala Glu	
119	455 460 465	
120	gca tca atc aac agc tcc gga aac agc tat act gaa atc cgg gcg cag	1536
121	Ala Ser Ile Asn Ser Ser Gly Asn Ser Tyr Thr Glu Ile Arg Ala Gln	
122	470 475 480	
123	ctc aat aac cgt tcg gga tgg ccg gca aag aaa acc gat caa ttg tct	1584
124	Leu Asn Asn Arg Ser Gly Trp Pro Ala Lys Lys Thr Asp Gln Leu Ser	
125	485 490 495	
126	ttc cgc tac tac gtt gac ttg acg gaa gct gta gaa gcg gga tat tcc	1632
127	Phe Arg Tyr Tyr Val Asp Leu Thr Glu Ala Val Glu Ala Gly Tyr Ser	
128	500 505 510 515	
129	gcc gaa gat ata aaa gtc aca gcc ggc tat aac gaa ggg gcc tcg gta	1680
130	Ala Glu Asp Ile Lys Val Thr Ala Gly Tyr Asn Glu Gly Ala Ser Val	
131	520 525 530	
132	tca gag ctg aag ccg cat gac gct tca aag cac att tac tat aca gaa	1728
133	Ser Glu Leu Lys Pro His Asp Ala Ser Lys His Ile Tyr Tyr Thr Glu	
134	535 540 545	
135	gtc agc ttc agc ggg gtt ttg att tat cca ggc ggt caa tcc gcc cat	1776
136	Val Ser Phe Ser Gly Val Leu Ile Tyr Pro Gly Gly Gln Ser Ala His	
137	550 555 560	
138	aaa aaa gaa gtg cag ttc cgc ctt tcg gca cca gac gga acg tct ttt	1824
139	Lys Lys Glu Val Gln Phe Arg Leu Ser Ala Pro Asp Gly Thr Ser Phe	
140	565 570 575	
141	tgg aac ccg gaa aat gac cac tct tat cag ggt ctg tca cat gcg ctt	1872
142	Trp Asn Pro Glu Asn Asp His Ser Tyr Gln Gly Leu Ser His Ala Leu	

RAW SEQUENCE LISTING

DATE: 02/25/2003

PATENT APPLICATION: US/09/576,778B

TIME: 11:45:52

Input Set : N:\CrF4\02242003\I576778C.raw

Output Set: N:\CRF4\02252003\I576778B.raw

```

143      580      585      590      595
144      ctg aag acg cgg tat att cct gtt tat gat gat gga cgg ctc gtt ttc      1920
145      Leu Lys Thr Arg Tyr Ile Pro Val Tyr Asp Asp Gly Arg Leu Val Phe
146      600      605      610
147      gga cat gag ccc ggt tac tag      1941
148      Gly His Glu Pro Gly Tyr
149      615
151 <210> SEQ ID NO: 2
152 <211> LENGTH: 646
153 <212> TYPE: PRI
154 <213> ORGANISM: Bacillus licheniformis
155 <400> SEQUENCE: 2
156      Met Lys Ala Leu Cys Leu Ala Leu Leu Val Ile Phe Ser Met Ser Ile
157      -25      -20      -15
158      Ala Ser Phe Ser Glu Lys Thr Arg Ala Ala Ser Ala Glu Glu Tyr Pro
159      -10      -5      -1 1
160      His Asn Tyr Ala Glu Leu Leu Gln Lys Ser Leu Leu Phe Tyr Glu Ala
161      5      10      15
162      Gln Arg Ser Gly Arg Leu Pro Glu Asn Ser Arg Leu Asn Trp Arg Gly
163      20      25      30      35
164      Asp Ser Gly Leu Glu Asp Gly Lys Asp Val Gly Leu Asp Leu Thr Gly
165      40      45      50
166      Gly Trp Tyr Asp Ala Gly Asp His Val Lys Phe Gly Leu Pro Met Ala
167      55      60      65
168      Tyr Ser Ala Ala Ile Leu Ser Trp Ser Val Tyr Glu Tyr Arg Asp Ala
169      70      75      80
170      Tyr Lys Glu Ser Gly Gln Leu Asp Ala Ala Leu Asp Asn Ile Lys Trp
171      85      90      95
172      Ala Thr Asp Tyr Phe Leu Lys Ala His Thr Ala Pro Tyr Glu Leu Trp
173      100      105      110      115
174      Gly Gln Val Gly Asn Gly Ala Leu Asp His Ala Trp Trp Gly Pro Ala
175      120      125      130
176      Glu Val Met Pro Met Lys Arg Pro Ala Tyr Lys Ile Asp Ala Gly Cys
177      135      140      145
178      Pro Gly Ser Asp Leu Ala Gly Gly Thr Ala Ala Ala Leu Ala Ser Ala
179      150      155      160
180      Ser Ile Ile Phe Lys Pro Thr Asp Ser Ser Tyr Ser Glu Lys Leu Leu
181      165      170      175
182      Ala His Ala Lys Gln Leu Tyr Asp Phe Ala Asp Arg Tyr Arg Gly Lys
183      180      185      190      195
184      Tyr Ser Asp Cys Ile Thr Asp Ala Gln Gln Tyr Tyr Asn Ser Trp Ser
185      200      205      210
186      Gly Tyr Lys Asp Glu Leu Thr Trp Gly Ala Val Trp Leu Tyr Leu Ala
187      215      220      225
188      Thr Glu Glu Gln Gln Tyr Leu Asp Lys Ala Leu Ala Ser Val Ser Asp
189      230      235      240
190      Trp Gly Asp Pro Ala Asn Trp Pro Tyr Arg Trp Thr Leu Ser Trp Asp
191      245      250      255
192      Asp Val Thr Tyr Gly Ala Gln Leu Leu Leu Ala Arg Leu Thr Asn Asp

```

RAW SEQUENCE LISTING

DATE: 02/25/2003

PATENT APPLICATION: US/09/576,778B

TIME: 11:45:52

Input Set : N:\CrF4\02242003\I576778C.raw

Output Set: N:\CRF4\02252003\I576778B.raw

```

193      260      265      270      275
194 Ser Arg Phe Val Lys Ser Val Glu Arg Asn Leu Asp Tyr Trp Ser Thr
195      280      285      290
196 Gly Tyr Ser His Asn Gly Ser Ile Glu Arg Ile Thr Tyr Thr Pro Gly
197      295      300      305
198 Gly Leu Ala Trp Leu Glu Gln Trp Gly Ser Leu Arg Tyr Ala Ser Asn
199      310      315      320
200 Ala Ala Phe Leu Ala Phe Val Tyr Ser Asp Trp Val Asp Thr Glu Lys
201      325      330      335
202 Ala Lys Arg Tyr Arg Asp Phe Ala Val Arg Gln Thr Glu Tyr Met Leu
203      340      345      350      355
204 Gly Asp Asn Pro Gln Arg Ser Phe Val Val Gly Tyr Gly Lys Asn
205      360      365      370
206 Pro Pro Lys His Pro His His Arg Thr Ala His Gly Ser Trp Ala Asn
207      375      380      385
208 Gln Met Asn Val Pro Glu Asn His Arg His Thr Leu Tyr Gly Ala Leu
209      390      395      400
210 Val Gly Gly Pro Gly Arg Asp Asp Ser Tyr Arg Asp Asp Ile Thr Asp
211      405      410      415
212 Tyr Ala Ser Asn Glu Val Ala Ile Asp Tyr Asn Ala Ala Phe Thr Gly
213      420      425      430      435
214 Asn Val Ala Lys Met Phe Gln Leu Phe Gly Lys Gly His Val Pro Leu
215      440      445      450
216 Pro Asp Phe Pro Glu Lys Glu Thr Pro Glu Asp Glu Tyr Phe Ala Glu
217      455      460      465
218 Ala Ser Ile Asn Ser Ser Gly Asn Ser Tyr Thr Glu Ile Arg Ala Gln
219      470      475      480
220 Leu Asn Asn Arg Ser Gly Trp Pro Ala Lys Lys Thr Asp Gln Leu Ser
221      485      490      495
222 Phe Arg Tyr Tyr Val Asp Leu Thr Glu Ala Val Glu Ala Gly Tyr Ser
223      500      505      510      515
224 Ala Glu Asp Ile Lys Val Thr Ala Gly Tyr Asn Glu Gly Ala Ser Val
225      520      525      530
226 Ser Glu Leu Lys Pro His Asp Ala Ser Lys His Ile Tyr Tyr Thr Glu
227      535      540      545
228 Val Ser Phe Ser Gly Val Leu Ile Tyr Pro Gly Gly Gln Ser Ala His
229      550      555      560
230 Lys Lys Glu Val Gln Phe Arg Leu Ser Ala Pro Asp Gly Thr Ser Phe
231      565      570      575
232 Trp Asn Pro Glu Asn Asp His Ser Tyr Gln Gly Leu Ser His Ala Leu
233      580      585      590      595
234 Leu Lys Thr Arg Tyr Ile Pro Val Tyr Asp Asp Gly Arg Leu Val Phe
235      600      605      610
236 Gly His Glu Pro Gly Tyr
237      615
239 <210> SEQ ID NO: 3
240 <211> LENGTH: 42
241 <212> TYPE: DNA
242 <213> ORGANISM: Artificial Sequence

```

VERIFICATION SUMMARY

DATE: 02/25/2003

PATENT APPLICATION: US/09/576,778B

TIME: 11:45:53

Input Set : N:\Crf4\02242003\I576778C.raw

Output Set: N:\CRF4\02252003\I576778B.raw

L:5 M:270 C: Current Application Number differs, Wrong Format
L:22 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:0
L:26 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:0
L:26 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:0